Development and Validation of a Needs-Assessment Instrument For Persons with Developmental Disabilities

**J. Michael Hennike,[1] Alexander M. Myers,[1,2]
Rodney E. Realon,[1] and Thomas J. Thompson[1]**

We conducted two studies to select an instrument that could be used statewide in North Carolina to assess level of intensity of support needs for all persons in the Developmental Disabilities Service System in a consistent, reliable, and valid manner. Study 1 compared three assessment instruments: two standardized instruments [the Inventory for Client and Agency Planning (ICAP) and the Developmental Disabilities Profile (DDP)] and the North Carolina Support Needs Assessment Profile[3] (NC-SNAP) in measuring the level or intensity of a person's needs. Study 2 reexamined the NC-SNAP after extensive revisions were made to it in an attempt to improve predictive validity. Results of this research indicated that the NC-SNAP is a reliable instrument capable of a valid assessment of levels or intensity of needs. North Carolina has adopted the NC-SNAP and its associated database program as the statewide assessment protocol. We discuss issues regarding evaluation of individual level of need and future research.

Faced with escalating Medicaid expenditures in the mid-1990s and with potential caps on growth and block grants for Medicaid at the

---

[1] Murdoch Center, Butner, North Carolina.

[2] To whom correspondence should be addressed at Murdoch Center P.O. Box 3000, 1600 C Street, Butner, North Carolina 27509

[3] Also known as the Developmental Disabilities Support Needs Assessment Profile or DD-SNAP.

201

Federal level, many states recognized the importance of determining the potential ramifications for citizens with developmental disabilities. North Carolina convened a blue ribbon panel consisting of representative stakeholders in the field, such as consumers and their families, public and private providers, advocates, and state and local officials. This diverse group quickly reached consensus on one issue: North Carolina had no consistent or standardized method to assess an individual's intensity of need for support and, as a result, it was impossible to determine whether individuals were being over-, under-, or well-served. Additionally, this contributed to absence of any cohesive statewide data on service need, which impeded state and local agencies and providers in their attempts to develop adequate resources and services. This was especially problematic as North Carolina had over 7300 persons on a waiting list for services, about 40% of whom were receiving no services at all. The blue ribbon panel established a Committee on Assessment and Service Array to identify or develop a reliable and valid assessment protocol that could assign "level of intensity of support needs" employing a 5-point hierarchical scale for individuals in North Carolina with developmental disabilities.

The research staff working with this committee initially reviewed efforts in place throughout the nation to meet similar missions. This review indicated that many states were struggling with the same issues and using a variety of methods to assess need. Despite extensive efforts by government agencies in numerous states, accurate assessment remained one of the most difficult and controversial issues facing human service organizations (Brown *et al.*, 1993). Ashbaugh (1996) reviewed assessment instruments in nine states. The two most widely used instruments at the time were the Inventory for Client and Agency Planning (ICAP) and the Developmental Disabilities Profile (DDP). A review of these instruments raised some concerns in terms of their intended use in North Carolina. The DDP has a scoring system that originally did not determine a level of need, although a leveling system was proposed by Myers and Stouffer (1993, Model 10). The ICAP is lengthy (77 adaptive items plus nine behavior items) and time-consuming to complete. In reviewing these instruments, the committee concluded that North Carolina required a more succinct and functional instrument, which would be easier to administer and score. Subsequently, we developed the North Carolina Support Needs Assessment Profile (NC-SNAP). We designed Study 1 to assess the efficacy of the NC-SNAP in assessing level of support needs for persons in North Carolina with developmental disabilities in relation to the other two commonly used instruments.

GENERAL METHOD

Instruments and Scoring Methods

The Inventory for Client and Agency Planning (ICAP)

The ICAP is a standardized assessment instrument developed by Bruininks *et al.* (1986). They developed the ICAP to assess status, adaptive functioning, maladaptive behavior, and service needs of clients. Combining the Adaptive Behavior Raw Score, and the General Maladaptive Index results in a numerical score that falls into one of nine levels. Descriptors are provided for odd-numbered levels, ranging from "Total personal care and intense supervision" (level 1) to "Infrequent or no assistance for daily living" (level 9). Because we were instructed to use a five-level system, we collapsed these nine levels into five using two alternate methods. Both methods involved collapsing the levels by pairing eight ICAP levels into four, leaving one level unpaired. "ICAP-1" created five levels by leaving level 1 unpaired and pairing the remaining levels (i.e., 1, 2–3, 4–5, 6–7, 8–9). "ICAP-9" refers to the alternate method that left level 9 unpaired (i.e., 1–2, 3–4, 5–6, 7–8, 9).

The Developmental Disabilities Profile (DDP)

The New York State Office of Mental Retardation and Developmental Disabilities developed the DDP in 1987 to simply and briefly document key characteristics of persons with developmental disabilities. Although the DDP was not developed as a leveling instrument, some researchers (cf., Myers and Stouffer, 1993) have attempted to convert scores into levels. After scoring the DDP, a factor is calculated for each of three domains (Adaptive, Maladaptive, and Medical) by dividing the individual's score in each domain by the highest score obtained by any individual in the sample for that domain. Totaling the three domain factors yields a final score ranging from 50 to 300. For the present investigation, we assigned scores to one of five levels [ranging from "Intensive needs" (level 1) to "Essentially independent" (level 5)] using the criteria developed by Myers and Stouffer (1993; Model 10). Because of North Carolina's plan for using an assessment instrument statewide, we reasoned that divisors obtained for each domain would approach the maximum as the sample increased to include the estimated 27,000 persons with developmental disabilities living in the state. We adopted two separate methods of calculating DDP scores: "DDPS" denotes

scores calculated using the divisor obtained as the highest score on a given domain from the sample in the study; "DDPM" refers to scores calculated using the highest score possible for each domain. This allows calculation for both the actual research sample and the theoretical maximum.

The North Carolina Support Needs Assessment Profile (NC-SNAP)

We developed the NC-SNAP in conjunction with the Murdoch Center Foundation under the guidance of the North Carolina Developmental Disabilities Advisory Work Group. The NC-SNAP directs the examiner to score 11 items on a 5-level scale in three support domains (daily living, health care, and behavioral). A score is then obtained for each domain by selecting the highest score marked for that domain. The overall score is equivalent to the highest domain score.

Independent Index of Support Need and Participant Selection

We next conceptualized descriptions of the five levels of support arrays, which would correspond to the five levels of need. Table I displays descriptors used for the support arrays and the five levels for each of the instruments used.

In the absence of a preexisting independent index of intensity of support need against which the instruments could be tested, we devised a method to create such a measure. First, we determined that participants would be selected for inclusion only if there was a consensus that they were currently receiving "good" (better than "acceptable") or "ideal" supports. The level of support provided to those persons receiving "good" or "ideal" supports could then be summarized and categorized into five "support intensity" levels, as described in Table I. The instruments were then administered. The instruments were considered valid if the level of need derived from the assessment tool matched that participant's assigned support intensity level. Inclusion of data from persons whose supports were not rated as "good" or better would likely yield an inaccurate Support Intensity level rating. Such inaccuracies would introduce error into the validity assessments, which are described later.

For both studies, surveys were sent to individuals with developmental disabilities (or their guardians), their current primary service provider, and to an independent case manager within five of North Carolina's 39 Area Mental Health Programs. The sampled Area Programs were selected by an independent committee as representative of North Carolina's demographic

Table I. Level of Intensity of Current Supports by Instrument

NC-SNAP level	Current supports	ICAP-1 level	ICAP-9 level	DDP level [M&S Model 10]
1	Essentially independent, may receive some periodic assistance	8–9	9	5 (scores between 50.00 and 74.99)
2	Receives some assistance on a routine basis, but not continuously. Requires some level of supervision but, again, not continuously	6–7	7–8	4 (scores between 75.00 and 99.99)
3	Receives continuous but not highly intensive assistance, but not 24-hr. awake staff	4 –5	5–6	3 (scores between 100.00 and 124.99)
4	Receives continuous and intensive assistance, with 24-h awake staff	2 –3	3–4	2 (scores between 125.00 and 149.99)
5	Receives continuous, intensive, and specialized assistance 24-h/day. For example, *Specialized (high need) Residential Settings, Skilled Nursing, Family Home* (Continuous & Specialized assistance), *Specialized Unit* (serving individuals with extreme needs)	1	1—2	1 (scores between 150.00 and 300.00)

range (i.e., a mixture of rural and urban, rich resources and poor, etc.). The survey consisted of a 5-point scale (1 = *poor*; 2 = *needs improvement*; 3 = *acceptable*; 4 = *good*; 5 = *ideal*) on which services were to be rated. Additionally, one of the three persons surveyed for each potential participant was asked to describe current supports. Participants were selected for inclusion in the study when all three parties responding to the survey indicated the individual was receiving "good" or "ideal" supports. In cases where there was no independent case manager, the survey was completed by the other two parties, and the Developmental Disabilities Coordinator for the relevant area program was polled using the survey instrument as to whether the individual's current supports were appropriate or not. The purpose of requiring agreement by independent raters was to reduce error that could result from biased reporting (e.g., case managers might be biased toward inaccurately reporting that their clients were receiving better than average services for fear that more honest reporting might have a negative impact on their job performance rating).

Each participant was assigned a "support intensity level" by comparing the information gathered from the listing of current supports (e.g.,

residential support, job coach, vocational services, educational program, etc.) to the descriptions found in the "Support Intensity" column of Table I. In Study 1, this assignment was done independently by three of the authors. Comparison of these independent assignments revealed nearly complete agreement (i.e., all the authors agreed on a participant's assigned support level) for 98.2% of the participants.

STUDY 1

Participants

The five Area Programs identified a total of 2332 potential participants. Of these, 553 were selected for inclusion in this study, using the "survey consensus" criterion described earlier.

Administration of the Assessment Instruments

We used a standardized training protocol to teach the participants' case managers how to complete and score each of the assessment instruments. All training was videotaped for review by an independent Investigation Oversight Committee to ensure the protocol was followed properly and that procedures for each instrument were taught in a neutral, unbiased manner.

Each case manager received a binder containing the assessment instruments for the participant(s) for whom he or she was responsible. We inserted the instruments into binders in random order and instructed the case managers to complete them in the order of presentation. In addition, each binder contained a time log for recording the time spent completing each instrument, and a satisfaction survey to report the case manager's subjective impressions on the ease, comprehensiveness, and preference for each instrument.

Upon return of the assessment binders, research assistants checked all instruments to ensure that all items were completed. The assistants also checked the accuracy of the calculation of the final score on the NC-SNAP and scored each instrument (DDP and ICAP) using software that had been specifically developed for each instrument. We entered the data into a master Excel database, and subsequently converted it to an SPSS data file. An independent consultant hired on contract from a local university completed the statistical analyses using SPSS 11.5. The consultant completed analyses to examine both the reliabilities and validities of the various instruments and their scoring systems.

Assessment of Reliability

We reassessed ten percent of the original sample to obtain measures of interrater reliability for each of the assessment instruments. A research assistant completed this second set of instruments by interviewing another person who the case manager had identified as being very familiar with the participant. We compared this second set of ratings to those generated by the case manager, using two indices of reliability: Pearson's *r*, and intraclass correlation coefficients (ICCs). Pearson's *r* assesses the degree to which the two independent ratings rank the participants in the same order, but it is not sensitive to differences in anchor points. Therefore, we also looked at intraclass correlation coefficients assuming both participants and raters were random, looking at "absolute" agreement. We report the ICCs for both individual and average ratings. We also calculated "percent agreement" for both exact matches and matches within (plus or minus) one support level.

Assessment of Validity

In order to assess the validity of the scores generated by the various instruments, we compared these scores to the support intensity levels, which we derived as described previously. For each instrument and/or scoring system, we cross-tabulated the actual levels with the scores derived from administering the instruments. In addition to presenting the tabulated data, we also computed Pearson's *r*, Cohen's kappa coefficients, percent exact match, match within one level, and percent over/under estimation.

Results

Reliability

Interrater reliability results for Study 1 are presented in Table II. The correlations indicate acceptable interrater reliability for all three instruments using the various leveling methods. The highest correlations were observed for the ICAP-1 and the ICAP-9. The percentage of exact level match was highest for the ICAP-9 and the NC-SNAP. The match rate within (plus or minus) one level reached 100% for the ICAP-1, and nearly 90% for the ICAP-9, and the NC-SNAP.

Validity

ICAP. Consistent with the leveling systems described previously, the levels identified for the ICAP and the DDP progress from most need (level

Table II.   Study 1 Reliability Measures: Correlations and Percent Match Between the Two Administrations for Each Instrument ($n = 58$)

Instrument	Pearson's r	ICC$_{(1)}$	ICC$_{(average)}$	% Match	% $\pm$ 1
DDPM	.694	.697	.822	44.83	87.93
DDPS	.764	.762	.865	50.00	82.76
ICAP-1	.846	.846	.917	58.62	100.00
ICAP-9	.835	.835	.910	70.69	96.55
NC-SNAP	.731	.731	.844	70.69	96.55

1) to least need (level 5). As seen in Table III, the ICAP-9 and the ICAP-1 yielded scores that were positively correlated with support intensity level ($r = .305$ and $.316$, respectively). Cohen's kappas for the ICAP-9 and the ICAP-1 were .104 and .028, respectively. Using the commonly applied criterion of acceptable reliability of $\kappa > .70$, neither ICAP scoring method achieved adequate agreement with assigned support levels.

There was an exact match between ICAP-9 level and the assigned level for 31.3% of the cases. The match rate within one level (more or less) of the assigned level was 69.6%. The ICAP-9 underestimated need (i.e., the ICAP-9 level was lower than the assigned level) for 37.3% of the cases. It overestimated need (i.e., the ICAP-9 level was higher than the assigned

Table III.   Validity Results for the ICAP-9 and the ICAP-1

	Support intensity level				
	1	2	3	4	5
ICAP-9[a]					
1	101	19	10	26	0
2	90	20	17	53	3
3	55	26	26	42	2
4	8	6	13	24	2
5	1	1	3	3	2
ICAP-1[b]					
1	38	5	5	6	0
2	108	19	12	38	1
3	78	34	21	59	4
4	28	11	24	30	1
5	3	3	7	15	3

Note. 1 represents *extreme need* and 5 represents *little need* ($n = 553$).
[a]Pearson's $r = .305$; Cohen's kappa $= .104$; Exact match $= 31.28\%$; Within $\pm 1 = 69.62\%$; Underestimate need $= 37.25\%$; Overestimate need $= 31.46\%$.
[b]Pearson's $r = .316$; Cohen's kappa $= .028$; Exact match $= 20.07\%$; Within $\pm 1 = 66.72\%$; Underestimate need $= 56.24\%$; Overestimate need $= 23.69\%$.

Table IV. Validity Results for the DDPM and the DDPS

	Support Intensity level				
	1	2	3	4	5
DDPM[a]					
1	13	3	4	3	0
2	63	15	5	13	0
3	81	16	8	30	1
4	54	22	20	32	1
5	44	16	32	70	7
DDPS[b]					
1	76	19	9	18	0
2	64	16	3	16	0
3	58	11	15	18	1
4	25	20	19	33	1
5	32	6	23	63	7

Note. 1 represents *extreme need* and 5 represents *little need* ($n = 553$).
[a] Pearson's $r = .328$; Cohen's kappa $= .000$; Exact match $= 13.56\%$; Within $\pm 1 = 51.18\%$; Underestimate need $= 75.59\%$; Overestimate need $= 10.85\%$.
[b] Pearson's $r = .395$; Cohen's kappa $= .083$; Exact match $= 26.58\%$; Within $\pm 1 = 62.39\%$; Underestimate need $= 58.05\%$; Overestimate need $= 15.37\%$.

level) for 31.5% of the cases. The ICAP-1 yielded an exact match with the assigned level for 20.1% of the cases. The match rate within one level (more or less) than the assigned level was 66.7%. The ICAP-1 underestimated need for 56.2% of the cases. It overestimated for 23.7% of the cases.

DDP. As seen in Table IV, the DDPM and the DDPS yielded levels that were positively correlated with assigned support intensity level ($r = .328$ and $.395$, respectively). These correlations were slightly higher than those found for the ICAP. Cohen's kappa scores were lower than those found for the ICAP ($\kappa = .000$ and $.083$, respectively). The Cohen's kappa scores indicate that neither DDP leveling method achieved adequate agreement with the assigned support levels.

There was an exact match between the DDPM level and the assigned level for 13.6% of the cases. The match rate within one level was 51.2%. Need was underestimated for 75.6% of the cases. It overestimated need for 10.9% of the cases. The exact match between the DDPS level and assigned level was 26.6%. Within one level, the match rate was 62.4%. Need was underestimated for 58.1% of the cases. It overestimated need for 15.4% of the cases.

NC-SNAP. The NC-SNAP leveling system progresses from least need (level 1) to most need (level 5). As seen in Table V, the NC-SNAP levels were positively correlated with the assigned level ($r = .274$). This correlation was lower than those found for the ICAP and the DDP. The Cohen's

Table V.   Validity Results for the NC-SNAP

| | Support intensity level | | | | |
	1	2	3	4	5
SNAP					
1	0	2	1	4	2
2	0	2	23	58	65
3	0	2	7	27	33
4	0	0	6	29	37
5	0	0	7	77	171

Note. Pearson's $r = .274$; Cohen's kappa $= .091$; Exact match $= 37.79\%$; Within $\pm 1 = 69.26\%$; Underestimate need $= 16.64\%$; Overestimate need $= 45.57\%$. 1 represents *extreme need* and 5 represents *little need* ($n = 553$).

kappa score ($\kappa = .091$) exceeded all others, except that found for the ICAP-9. Nevertheless, the κ score indicates that the NC-SNAP leveling method did not achieve adequate agreement with assigned support levels.

The exact match between the NC-SNAP level and the assigned level was 37.8%, the highest for the five leveling methods. The match rate within one level was 69.3%. The NC-SNAP underestimated need for 16.6% of the cases. It overestimated need for 45.6% of the cases.

Completion Time and Preference Measures. The NC-SNAP was the quickest instrument to complete, with a mean duration of 15 min, and a range of 2–45 min. The DDP averaged 22.5 min to complete, with a range of 4 to 120 min. The ICAP was the longest to complete, averaging 41 min, with a range of 5–120 min. The time needed to complete the scoring was not studied. The NC-SNAP is self-scoring, whereas both the DDP and ICAP scoring can be facilitated using computer software. The examiners showed no significant differences in their selection of a *most preferred* instrument. They ranked the NC-SNAP as the *least preferred* significantly more often (53.5%, as opposed to 25.9% for the ICAP and 21.1% for the DDP).

Discussion

Although all of the instruments demonstrated a positive correlation between the attained level and the assigned level, these correlations were low. The percent of exact match exceeded 30% only for the ICAP-9 (31.3%) and the NC-SNAP (37.8%). Kappa scores for all of the instruments were insufficient to indicate validity according to the conventional standard. Both the ICAP and the DDP leveling systems tended to underestimate need, whereas the NC-SNAP tended to overestimate need. The NC-SNAP was

easy to use, fast, and self-scored. The ICAP was longer and more complicated to score. The ICAP and NC-SNAP both received positive ratings from users, although more users ranked the NC-SNAP lowest in terms of preference. Anecdotal reports indicated that users preferred the ICAP for its perceived comprehensiveness whereas they liked the NC-SNAP for its ease of administration.

In reviewing the relative merits of the instruments following completion of Study 1, the authors and the North Carolina Developmental Disabilities Section recognized that none produced predictive values relative to the assigned support intensity level sufficient to warrant its use as a statewide assessment tool. The ICAP is a copyrighted instrument we could not amend in order to enhance its predictive properties. Therefore, we undertook an item-by-item re-analysis of the NC-SNAP using the original Study 1 database. That is, we used the data from Study 1 to examine each item of the NC-SNAP to determine whether level revision or deletion of items would improve the instrument's predictive ability. Recognizing that further changes could be made that would not be properly analyzed using the previous database, we made changes and designed a second study to determine whether the revised NC-SNAP would yield greater concurrent validity.

STUDY 2

Methods

Participants

Using the same participant selection criterion used in Study 1, we identified 175 persons from a population of 595 persons with developmental disabilities at a sixth Area Program as potential participants. In Study 2, however, we attempted to stratify the sample on the basis of level of support intensity by identifying 20 participants at each support intensity level. In actuality, we selected 80 participants, as only 13 participants were available for support intensity level 1, 12 for level 2, and 13 for level 5.

Administration of the NC-SNAP

There were several key procedural differences in Study 2. First, as our intent in the second study was to determine the predictive ability of a revised NC-SNAP, we did not administer either the DDP or the ICAP.

Second, instead of teaching case managers to complete the instrument as we did in the first study, two of the researchers completed the NC-SNAP during interviews with the participants' case managers. Because there were only 80 participants, this process was much more efficient than training multiple case managers, some of whom would need to assess only one participant. Third, for each participant, a third researcher independently assigned the support intensity level corresponding to current supports and services. We did not reveal the participants' assigned support intensity level to the two researchers administering the NC-SNAP until the instrument was fully scored.

A secondary purpose of Study 2 was to further refine the NC-SNAP, if necessary, through an error analysis of cases where the instrument's score did not match the participant's assigned support intensity level. Therefore, following the completed scoring of the NC-SNAP, while still in the presence of the participant's case manager, the researcher/examiner opened a sealed envelope containing the participant's assigned support intensity level. If the NC-SNAP and support intensity levels did not match precisely, the researcher and case manager discussed variables that differed in an attempt to determine the source of error. This allowed us to identify items on the NC-SNAP which might require modification and identify errors in the support intensity level assignment (e.g., the third researcher described above may have made an incorrect support intensity level assignment, the information provided about the participant's current supports and services could have been incomplete or could have been outdated due to changes in the participant's needs). By correcting errors in support intensity assignment we were able to determine whether the NC-SNAP had actually erred in its prediction of level of support need.

Assessment of Reliability

We assessed the reliability of the revised NC-SNAP in several different ways. To obtain a measure of test-retest reliability, we asked certified NC-SNAP examiners at Murdoch Center, a large state residential facility for persons with mental retardation, to repeat the administration of an initial NC-SNAP for 103 individuals who had been originally assessed during the previous month. A researcher (unaware of the scores obtained on the previous administrations) provided data for assessing interrater reliability by completing a third NC-SNAP, using an interview format with the original examiner, for this same group of individuals shortly after the completion of the second administration. For both test–retest reliability (i.e., first and second administrations) and interrater reliability (i.e., second and third

administrations), we have computed Pearson's *r*, intraclass correlations for both individual and average scores, and percentage agreement.

Assessment of Validity

We assessed the validity of scores obtained on the revised NC-SNAP by comparing them to the support intensity levels, which were derived as described previously. We cross-tabulated the actual levels with the scores derived from administering the revised NC-SNAP, and for the post hoc analysis of corrected support intensity levels. In addition to presenting the tabulated data, we also computed Pearson's *r*, Cohen's kappa coefficients, percent exact match, match within one level, and percent over/under estimation.

Results

Reliability

Test–retest reliability results are presented in Table VI. Both Pearson and intraclass correlations were high for individual domain scores, and for the overall NC-SNAP score. There was a 96% match rate for the overall NC-SNAP score. Match rates for Daily Living, Health Care, and Behavioral domains were 97.0, 87.1, and 83.2% respectively. Match rates within one level (more or less) were 100, 94, and 96% respectively.

Interrater reliability results for Study 2 are presented in Table VII. Both Pearson and intraclass correlations were high for individual domain scores and for the overall NC-SNAP score. There was a 92.1% match rate for the overall NC-SNAP score. Match rates for the Daily Living, Health Care and Behavioral domains were 95.0, 77.2, and 73.3%, respectively. Match rates within one level were 100, 90.1, and 99.0%, respectively.

Validity

The results of the revised NC-SNAP are presented in Table VIII. There was a high positive correlation between the NC-SNAP level and the

Table VI. Test–Retest Reliability Results From Scoring the Revised NC-SNAP Including Pearson's *r*, $ICC_{(1)}$, $ICC_{(average)}$, and Percentages of Agreement ($n = 101$)

Scale	Pearson's *r*	$ICC_{(1)}$	$ICC_{(average)}$	Match (%)	±1 (%)
Daily living	.920	.915	.956	97.03	100.00
Health care	.926	.927	.962	87.13	94.06
Behavioral supports	.821	.813	.897	83.17	96.04
Overall NC-SNAP	.918	.917	.957	96.04	100.00

Table VII.   Interrater Reliability Results From Scoring the Revised NC-SNAP Including Pearson's r, ICC$_{(1)}$, ICC$_{(average)}$, and percentages of agreement ($n = 101$)

Scale	Pearson's r	ICC$_{(1)}$	ICC$_{(average)}$	Match (%)	±1 (%)
Daily living	.882	.879	.936	95.04	100.00
Health care	.881	.874	.933	77.23	90.10
Behavioral supports	.836	.828	.906	73.27	99.01
Overall NC-SNAP	.856	.854	.924	92.08	100.00

assigned support intensity ($r = .723$). Cohen's kappa ($\kappa = .614$) approached the common criterion for acceptable reliability ($\kappa = .70$). There was an exact match between the NC-SNAP level and the assigned support intensity of 70%, which nearly doubled the exact match rate for the previous version of the NC-SNAP. The match rate within one level of the assigned level was 86.3%. The NC-SNAP underestimated need for 6.3% of the cases. It overestimated need for 23.8% of the cases.

Following the administration of the NC-SNAP, where disagreements between the NC-SNAP level and the assigned support intensity were detected, we debriefed the case manager to determine if an error had been made in the original support intensity assessment. When an error was found, we corrected the support intensity level. We then repeated the previous analysis. Data from this analysis are found in Table IX.

Correction of the assigned support intensity errors resulted in a very high correlation between the NC-SNAP level and the support intensity level ($r = .956$). Cohen's Kappa ($k = .884$) exceeded the criteria for acceptable reliability. The exact match rate improved to 91.3%. The match rate within one level was 98.3%. With these corrections, the NC-SNAP underestimated need in 2.5% of the cases. It overestimated need in 6.3% of the cases. Our observation of the usefulness of correction procedures led to the adoption of quality assurance processes for the statewide application of the NC-SNAP, as described below.

Table VIII.   Validity Results for the Revised NC-SNAP

	SNAP				
	1	2	3	4	5
Support intensity level					
1	6	3	2	2	0
2	0	4	3	3	2
3	0	1	16	3	1
4	0	1	2	18	0
5	0	0	7	1	12

Note. Pearson's $r = .723$; Cohen's kappa $= .614$; Exact match $= 70.00\%$; Within $\pm 1 = 86.25\%$; Underestimate need $= 6.25\%$; Overestimate need $= 23.75\%$. 1 represents *little need* and 5 represents *extreme need* ($n = 80$).

Table IX. Validity Results for the Revised NC-SNAP Using Corrected Support Intensity Levels

	SNAP				
	1	2	3	4	5
Support intensity level					
1	6	1	1	0	0
2	0	7	1	0	0
3	0	1	20	2	0
4	0	0	1	25	0
5	0	0	0	0	15

Note. Pearson's $r = .956$; Cohen's kappa $= .884$; Exact match $= 91.25\%$; Within $\pm 1 = 98.25\%$; Underestimate need $= 2.5\%$; Overestimate need $= 6.25\%$. 1 represents *little need* and 5 represents *extreme need* ($n = 80$).

OVERALL DISCUSSION

None of the three instruments used in Study 1 was particularly effective at predicting the relative level of support needs for persons with developmental disabilities. However, the initial version of the NC-SNAP performed at least as well as the other two commonly-used instruments. Modifications made to the NC-SNAP on the basis of the findings in Study 1 produced an instrument that was found in Study 2 to have achieved our goal to provide a reliable and valid assessment in terms of level of acuity of need for persons with developmental disabilities. Secondary goals were also achieved: the NC-SNAP was found to be easy to administer, self-scoring, and reporting of scores straightforward.

There were two possible limitations to this investigation, both relating to the construct of "need." Clearly, an important component in assessing the accuracy of a measurement instrument is defining what is being measured. One limitation to this study was the absence of a preexisting independent index of need acuity against which the three instruments could be measured. We addressed this problem by identifying the five "Support Intensity" levels. Although we could not measure construct validity, approval of this model by the Committee on Assessment and Service Array provided face validity to our measurement strategy. Another potential limitation relating to the applicability of the NC-SNAP or any needs assessment instrument involves the definition of what constitutes an individual's "needs." The three categories of need measured by the NC-SNAP are typical of those used by public agencies to direct service delivery. The acuity of other "needs" (e.g., personal desires) cannot be measured with the NC-SNAP. In

the current environment of limited and sometimes reduced public spending, the NC-SNAP's focus on basic health and safety needs appears prudent.

Accurate identification of a person's need is a logical and necessary first step in the process of arranging supports and services. The NC-SNAP appears to aid in that need-determination process and does so in a parsimonious manner. Its use can facilitate the selection of appropriate supports, and its fourth page can serve as a template when beginning to develop a comprehensive support plan. The NC-SNAP may also help identify individuals who are currently under- or over-served, allowing reassignment of such persons to optimal supports. This should result in more efficient utilization of resources, in terms of improving providers' ability to allocate services and to develop future supports. Review of congregate data will allow responsive planning and development of services and supports. Ultimately, this could result in fiscal savings.

The purpose of the present investigation was to identify an assessment tool that would accurately predict level of support needs for persons being well served within North Carolina's Developmental Disabilities System. The NC-SNAP has been adopted for this purpose, along with a comprehensive database, which should permit analysis of trends both at the local and statewide level. There is an implementation protocol with built-in reliability checks and strategies to ensure accurate assessment.

The NC-SNAP is currently used to provide annual assessments for approximately 30,000 persons with developmental disabilities in North Carolina. A standardized competency-based format, which lasts approximately 3 h, has been used to train examiners across the state. Experienced Instructors certified by the authors conduct all training in order to assure consistency. Instructors provide periodic refresher training to examiners to maintain skills. Numerous quality assurance procedures are in place to monitor reliability and validity. These include "look behind" assessments of reliability by instructors who provide corrective feedback to examiners when errors are noted. Scores are reported to a centralized database where a coordinator selects unusual patterns to review for errors. We believe that this "research to practice" model has maintained the efficacy of the assessment instrument as it has been adopted in widespread use.

Trials are underway to determine whether NC-SNAP scores correlate with funding costs. Although several research designs are being considered to accomplish this task, there are serious obstacles when attempting to correlate acuity of need to financial expenses. First, in North Carolina an individual's supports are typically funded through multiple and complex sources. Attempts to track these funding "streams" (e.g., multiple local, state, and federal monies) can be surprisingly and frustratingly difficult. Second, persons receiving identical supports may expend vastly different

budgets because of the availability of natural supports. For instance, a person with extreme medical needs under the care of a dedicated parent in the family home will typically have associated costs that are far lower than a person with identical needs living in a skilled nursing setting. Third, service assignment can be a function of the availability of services; a person can receive a more or less expensive service than another individual simply due to the menu of services available locally. Fourth, there are persons who are under-served and persons who are over-served throughout the service system, rendering cost of supports invalid as a measure of need. Finally, cost of services may be affected by location and facility size, type, and ownership without differentiation between individuals with differing needs (Campbell and Heal, 1995). For instance, persons of differing levels of need residing in the same ICF/MR group home will receive the same Medicaid reimbursement.

Assessing the integrity of examiner training, performing reliability checks, and other quality assurance activities are essential to maintaining an accurate database. When used properly, the NC-SNAP can be an important component of a statewide service delivery system.

ACKNOWLEDGMENTS

REFERENCES

Ashbaugh, J. (1996). *Review of Assessment Tools and Procedures For Identifying Resources Needed By Persons With Developmental Disabilities* (Report prepared for Office of Health and Rehabilitation Services, Colorado Department of Human Services, Denver, CO). Human Services Research Institute, Cambridge, MA.

Brown, C. M., Hanley, A., Bontempo, A., and Manning, B. L. (1993). *The DDP and Resource Consumption,* New York State Office of Mental Retardation and Developmental Disabilities, Albany.

Bruininks, R. H., Hill, B., Weatherman, R. E., and Woodcock, R. (1986). *Inventory For Client and Agency Planning*, Riverside, Chicago, IL.

Campbell, E. M., and Heal, L. W. (1995). Prediction of cost, rates, and staffing by provider and client characteristics. *Am. J. Mental Retard.* 100: 17–35.

Myers and Stauffer Consultants (1993) *North Carolina ICF-MR reimbursement study*, Myers and Stauffer Consultants, Raleigh, NC (Final Report of Contract #MA-075–93 with the North Carolina Department of Medical Assistance).